CONFORMATIONAL ANALYSIS OF THE ACYL POCKET LOOP IN ACETYLCHOLINESTERASE COMPUTED BY MONTE CARLO METHODS WITH A GENERALIZED BORN MODEL OF SOLVATION

Louis Carlacci
Army High Performance Computing Research Center
Network Computing Services, Inc.
PO Box 581459, Minneapolis, Minnesota 55458-1459

Mark Olson and Charles B. Millard
Department of Cell Biology and Biochemistry
USAMRIID, 1425 Porter St., Fort Detrick, MD 21702-5011

ABSTRACT

Acetylcholinesterase catalyzes the hydrolysis of the neurotransmitter, called acetylcholine. Based on X-ray crystallography, the acyl pocket loop (APL; residues 287 to 290) was observed in two conformations, the native conformation (2ACE) and an alternative conformation (2DFP), which the APL interacts with an irreversible inhibitor. The free energy of the 2DFP state was 4 kcal/mol higher than that of the 2ACE state. The native state has a more favorable solvation free energy, due primarily to Arg 289 interactions. The information should be useful in the development of antidotes against biological and chemical warfare agents that target AChE.

INTRODUCTION

Acetylcholinesterase (AChE) catalyzes the hydrolysis of acetylcholine, a chemical messenger in cholinergic synapses, to prepare nerve cells and skeletal muscles for a subsequent nerve impulse. AChE is susceptible to rapid, stoichiometric, and irreversible inhibition by a class of organophosphorus acid anhydride (OP) inhibitors, for example, diisopropylphophorofluoridate (DFP) and phosphonate nerve agents. The acyl pocket loop (APL) of AChE is a polypeptide segment in the reactive site that interacts with the CH_3 group of the acyl moiety of acetylcholine. Millard et al. observed flexibility in the APL from the crystal structure of aged phosphonylated AChE. [1] The native conformation of the APL is that observed in the crystal structure without inhibitor bound (2ACE of the Protein Data Bank, PDB). [2,3] An alternative conformation was observed with the DFP inhibitor covalently bound to the catalytic serine residue (2DFP of the PDB). [1,3]

Our goal in this study was to calculate the structural and energetic properties of the APL. For this purpose, a robust conformational search approach was employed in the search for native and nonnative

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
01 JUL 2003	**N/A**	**-**

4. TITLE AND SUBTITLE	5a. CONTRACT NUMBER
Conformational Analysis Of The Acyl Pocket Loop In Acetylcholinesterase Computed By Monte Carlo Methods With A Generalized Born Model Of Solvation	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Army High Performance Computing Research Center Network Computing Services, Inc. PO Box 581459, Minneapolis, Minnesota 55458-1459**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release, distribution unlimited

13. SUPPLEMENTARY NOTES
See also ADM001523.

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			
unclassified	**unclassified**	**unclassified**	**UU**	**8**	

loop conformations. An accurate scoring potential was used to distinguish between the conformational states. The role of solvation should play an important role due to the fact that many crystallographic waters were found inside the reactive site gorge. [4]

The conformational search of the APL was computed using the molecular mechanics program developed and implemented by Carlacci. [5] The final conformations from multiple independent Monte Carlo simulated annealing (MCSA) runs starting from random loop conformations were used to generate a conformational library from which main chain (mc-) conformations were interactively separated based on their dihedral angle coordinates. [6] The configurational entropy of a cluster mc-conformations (or mc-conformer) was computed from a conformational library generated from room temperature random walk simulations (or Metropolis Monte Carlo; MMC). [7]

Recent work by Qiu et. al [8] and Dominy & Brooks [9] should increase the use of implicit solvent models in molecular mechanics calculations. They developed a fast method for calculating approximate Born radii for use in the Generalized Born (GB) model of solvent polarization. The GB/SA model of solvation by Qiu is the sum of solute-solvent electrostatic polarization computed by GB method and solvent cavitation and solute-solvent van der Waals terms evaluated together by the method of solvent accessible surface (SAS).

The goals of this study were to predict and characterize experimentally observed APL conformations and other, low energy, alternative states. We wanted to know the difference in energy between the states, the structural relationships, and the role played by solvation and entropy.

MATERIALS AND METHODS

Heavy atom Cartesian coordinates of AChE defined as 2ACE [1] in the PDB [3] were used as starting points for the Monte Carlo calculations. In the construction of coordinates of the full protein, a missing loop, several missing side chains and hydrogen atoms were added. The geometry of the full protein was optimized by low temperature MCSA. The IUPAC nomenclature for amino acids was employed. [10]

COORDINATES

The geometry of the protein was generated in the ECEPP/2 reference frame, which the bond lengths and bond angles are fixed, and dihedral angles may be varied. [11] L-amino acids were employed. Ionizable side chains were put in the ionized state. The Cartesian coordinates of segments of the simulation volume that moved were generated from dihedral angles and Rigid body coordinates. [6]

ENERGY

The total energy (F) was the sum of the conformational energy of the simulation volume (E_{local}), free energy of hydrophobicity (G_{SA}), the GB solvent polarization energy (G_{GB}), and a weighted polypeptide chain loop closing constraint ($W \cdot F_{pep}$). The total energy was given by

$$F = E_{local} + G_{SA} + G_{GB} + W \cdot F_{pep} \qquad \text{(Eq. 1)}$$

The conformational energy was the sum of non-bonded, electrostatic and torsion potentials [11] computed with the OPLS united atom (UA) parameter set [12]. A vacuum dielectric constant (ε=1) was used. The local region energy was the sum of all intra-segment and inter-segment energies of segments that moved plus the sum of all interaction energies of these segments with the rest of the simulation volume.

The free energy of hydrophobicity was based on SAS computed using MSEED [13] with a 1.4 Å probe radius. Solvation parameters of hydrocarbons by Qiu et al. were used. [8] The GB solvent polarization energy was computed by the method developed by Qui et al. [8] OPLS UA charges [12] were employed. In this treatment, solvation free energy was the sum of GB solvent polarization energy of polar atoms and free energy of energy of hydrophobicity of hydrocarbons computed by SAS.

The polypeptide chain loop closing constraint forces the ends of the loop to dock against the protein. The loop closing weight (W in Eq. 2) was gradually introduced in the beginning of an MCSA run. This facilitated the removal of bad contacts and improves conformational sampling.

X-RAY STRUCTURE REFINEMENT

Initial dihedral angles for the APL calculation were generated by optimizing the overlay of the heavy atoms of the computed coordinates onto the X-ray coordinates. Energy minimization starting from the most populated conformations was used to optimize the coordinates of missing atoms. Energy minimizations employed solvation free energy based on SAS [6] and amber all atom parameters [14]. A missing loop, residues 485 to 489, was added. Final geometry optimizations consisted of multiple copy, low temperature MCSA runs, that used OPLS UA parameters and the GB/SA model of solvation.

SIMULATION VOLUME

The simulation volume (or the part of the protein included in energy calculations) consisted of the segments of the protein that were allowed to move during the calculations (the computed segments and select surrounding local region (SLR) side chains that interact with the computed segments} and SLR segments that interact with the movable parts but remain fixed during the calculation. The simulation volume employed included all residues that might interact with the APL (residues that are within approximately 12 angstroms from the APL). For various stages of the calculation, the segments and residues of the simulation volume are given in Table 1.

Table 1. Acyl pocket loop simulation volume at various stages of the calculation. [a]

stage [b]	CS [c]	SLR [d]	SLR-SC [e]
MCSA I	287 to 289 Ile-Phe-Arg-Phe	residues within 12 Å of CS	Tyr 121, Trp 279, Leu 282, Phe 284, Ser 286, Phe 330, Phe 331
MMC I	287 to 289	residues within 12 Å of CS	same as above
MMC II & MCSA II	CS + SLR of stage MCSA I	residues within 12 Å of CS	none

a) The simulation volume consisted of computed segments and SLR residues. b) Stages of the calculation, MCSA I: multiple independent MCSA prediction runs starting from random conformations,
MMC I: MMC runs starting from lowest energy conformations (LEC's) of predicted states,
MMC II: multiple copy MMC runs starting from the LEC's of MMC I.
MCSA II: multiple copy low temperature MCSA starting from LEC's of MMC II.
In MMC II and MCSA II, segments of the simulation volume of the prediction stage were refined.
c) Computed segment residue number and 3-letter code from the PDB. [3] d) Surrounding local region.
e) For SLR side chains that move, residue number and 3-letter code from the PDB.

MONTE CARLO METHODS

An independent MCSA loop prediction run started out with a high Boltzmann temperature factor (RT in kcal/mol; where R was the gas constant, and T was the absolute temperature in Kelvin) which facilitated the removal of bad contacts in the initial random conformation and the adoption of the native fold and other lower energy states. [5] RT was gradually lowered to room temperature. Trial conformations were accepted or rejected according to the Metropolis criteria. [15] Separate values of RT were used in the evaluation of main chain and side chain trial conformations. Table 2 summarizes Monte Carlo parameters for MCSA prediction runs. Room temperature (300 Kelvin) corresponds to RT=0.6 kcal mol^{-1}.

Table 2. Summary of parameters for MCSA prediction runs

parameter [a]	equilibration [b]	prediction [c]
RT main chain:	7.0 to 4.0	4.0 to 0.6
RT side chain:	5.0 to 2.0	2.0 to 0.6
NSTEP	15	10
NCON	4,000 (10,000 in step 1)	4,000

a) RT: Boltzmann temperature factor in kcal/mol. NSTEP: the number of steps. NCON: the number of trial conformations evaluated in a step. b) Parameters of equilibration.c) Parameters of prediction.

The final conformations of independent MCSA runs were interactively separated into mc-conformers. For each mc-conformer, the LEC was used as a starting point for MMC I runs. The LEC of each MMC I run was the starting point for MMC II runs, which employed a larger simulation volume and number of computed segments. The conformations and energies saved during the MMC II calculations were used in the evaluation of the thermodynamic properties. Low temperature MCSA II geometry optimization runs started from the LEC's of the MMC II calculations. Protocol parameters for MMC and low temperature MCSA optimizations are given in Table 3.

Table 3. Summary of MMC and low temperature MCSA parameters.[a]

parameter [b]	MMC pre-equil	MMC equilibration	MMC production	MCSA pre-equil	MCSA equilibration + production
RT	0.7 to 0.6	0.6	0.6	0.6 to 0.2	0.2
NCON					
MMC I	10K	620K × 1 PE	-	-	-
MMC II	10K	50K × 1 PE	15K × 4 PE	-	-
MMC II' [c]	-	18K × 1 PE	12K × 4 PE	-	-
MCSA II'	-	-	-	10K × 8 PE	40K × 8 PE

a) Stages of the calculation are pre-equilibration (pre-equil), equilibration and production. Note, 1 PE is one processor and a multiple parallel processor machine. b) In each stage, NSTEP was one; And, NCON trial conformations were evaluated. b) In the MMC II' stage, the starting conformation was the conformation of the LEC of stage MMC II with the side chain of Trp 279 manually adjusted and optimized to assume the X-ray conformation.

CONFORMER SEPARATION

A mc-conformer of a loop was a group of conformations with similar values of main chain dihedral angles. A main chain / side chain conformer was a group of conformations with similar values of

backbone and side chain dihedral angles. Clusters of mc-conformations were interactively separated from residue ϕ/ψ-scatter plots. [6] For each conformation of a mc-conformer, an alphanumeric tag that specified the conformations of the ϕ/ψ-pair and ω (called a mc-rotamer tag) was assigned to the residues A mc-conformer was a group of conformations with the same sequence of mc-rotamer tags.

RESULTS AND DISCUSSION

Based on the accessibility of the energy basin in MCSA prediction runs, entropy strongly favors the DFP-like state. Starting from random conformations, the 2DFP-like basin was found 117 times (states ranked 1 & 2 in Table 4). The 2ACE-like basin was found only 4 times (state ranked 4 in Table 4). The inaccessibility of the 2ACE basin is probably due to the loss of entropy associated with inserting the Phe 288 side chain between those of Phe 331 and Trp 233, and the lack of flexibility of other segments of the SLR. Note, only select side chains of the SLR were allowed to move at this stage. The last five columns in Table 4 give backbone (bb) and side chains (sc) root mean square deviations (rmsd's) of APL residues after computed LEC's were superimposed on the X-ray structures, and the computed LEC's were superimposed on the conformations of the conformer (LEC/mer). The LEC/mer result is a measure of cluster compactness.

Table 4. Characterization of 2638 final APL conformations from independent MCSA prediction runs.

rank [a]	mc-type [b]	ncon [c]	ΔE [d]	mc-rotamer tags[e]				LEC ACE bb [f]	LEC ACE sc [g]	LEC DFP bb [h]	LEC DFP sc [i]	LEC mer bb [j]
				Ile	Phe	Arg	Phe					
1	out	49	0.0	A	B2	F	B1	2.7	7.0	2.3	4.6	0.4
2	out	68	0.8	A	B2	F	B2	2.3	6.7	1.9	5.2	0.5
3	out	5	8.6	A	F	B2	B4	1.2	4.7	3.1	6.1	0.2
4	ext	4	9.9	B2	B2	F	B1	3.1	7.8	2.8	6.8	0.4
5	in	4	10.3	B3	L	B3	B3	0.4	1.7	3.4	6.3	0.3
6	out	3	11.3	A	F	B1	B4	1.3	3.1	3.0	5.3	0.2
7	in	22	11.4	B4	L	B1	B1	1.9	5.7	2.3	2.9	0.4
8	-	24	12.6	A	B4	L	B4	1.7	2.1	3.6	6.5	0.3

a) Conformer rank in order of increasing energy. b) Main chain structure type; "in": 2ACE-like, "out": 2DFP-like, "ext": beta-strand. c) Number of conformations in conformer. d) Relative total energy of the LEC in kcal/mol. e) mc-rotamer tags for each residue of the APL from N- to C-termini. A: α-helix region, B: β-strand region (B1: right-handed twisted, B2: left-handed twisted), F: turn region (centered on phi=-90, psi=10), L: left-handed helix region. f) backbone rmsd between the LEC and 2ACE X-ray conformation. g) side chain rmsd between the LEC and 2ACE X-ray conformation. h) backbone rmsd between the LEC and the 2DFP X-ray conformation. i) side chain rmsd between the LEC and the 2DFP X-ray conformation. j) backbone rmsd between the LEC and conformations of the conformer.

The free energy is the total energy at the maximum of the population versus energy histogram plus the configurational entropy contribution (Figure 1). The free energies of states in Table 5 with an asterisk appended to the rank were obtained from MMC runs after the side chain of Trp 279 was put in the native conformation, which is inside the gorge, and optimized. The interaction of Trp 279 side chain with a buried water molecule that was not explicitly included in these calculations was probably inaccurately represented by the GB model.

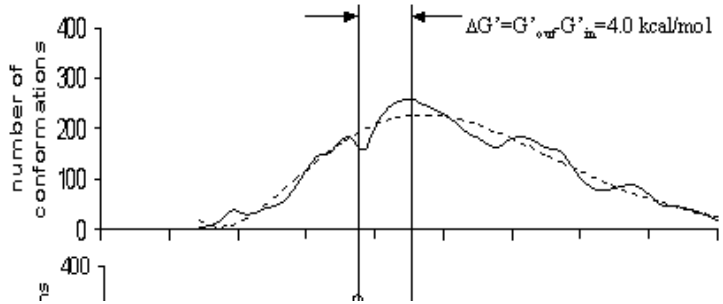

Figure 1. Free energy histogram. The free energy, ΔG', which does not include the configurational entropy, is the total energy at the maximum of the population versus energy histogram. The step size of the histogram is 1.2 kcal/mol.

The free energy of the 2ACE-like state (ranked 1 in Table 5) is 4 kcal lower than the next state, which is the 2DFP-like state (ranked 2). Solvation free energy favors the 2ACE-like state by ΔG_{sol}=-69.8kcal/molversus the 2DFP-like state. Internal conformational energy favors the alternative states by ΔE_{int}=-63.3 kcal/mol versus the 2ACE-like state. The data supports the proposition that the dielectric environment governs the transition between the native and alternative states.

Table 5. Energy properties of APL conformers obtained from MMC II calculation. [a]

	mc-					mc-rotamer tags [h]			
rank [c]	type [b]	ΔG [d]	S_{mic} [e]	ΔE_{int} [f]	ΔG_{sol} [g]	Ile	Phe	Arg	Phe
1*	in	0.0	4.2	0.0	0.0	B3	L	B3	B3
2*	out	4.0	4.8	-63.3	69.8	A	B2	F	B2
3*	ext	10.4	6.2	-109.6	120.6	B2	B2	F	B1
4*	out	13.4	5.8	-57.6	77.5	A	B2	F	B1
5*	in	16.0	4.2	-100.8	110.4	B4	L	B1	B1
6	out	21.0	2.6	60.4	-37.5	A	F	B1	B4
7	ext	22.0	5.4	-77.9	106.2	B2	B2	F	B1
8	-	22.1	3.7	88.2	-59.1	A	B4	L	B4

a) Thermodynamic properties of 4000 conformations are analyzed. b) Main chain structure type. c) Rank of conformer. An asterisk indicates that the starting conformation of the MMC run had Trp 279 adjusted to the X-ray conformation. d) Free energy at maximum of population versus energy histogram plus entropy contribution. e) Configuration entropy in cal/(K mol). [7] f) Relative conformational energy of the lowest energy conformation. g) Relative solvation free energy of the lowest energy conformation. h) See footnote e) mc-rotamer tags for each residue of the APL from N- to C-termini. A: α-helix region, B: β-strand region (B1: right-handed twisted, B2: left-handed twisted), F: turn region (centered on phi=-90, psi=10), L: left-handed helix region.

The native and alternative states were stabilized by different types of interactions with Arg 289. Table 6 summarizes the residue free energy contributions due to inter-residue conformational energy and inter-residue GB polarization energy. The free energy components of Arg 289 were significantly different between the native and alternative states. We found that solvation free energy of Arg 289

favored the native state over the alternative states by over 137 kcal/mol. And, inter-residue electrostatic and non-bonded energies favored the alternative states by over 128 kcal/mol. The residue free energy is the sum of intra-residue and inter-residue non-bonded and electrostatic energies, intra-residue and inter-residue GB solvent polarization energies, and free energy of hydrophobicity.

Table 6. Free energies components of individual residues of the APL. [a]

	$E_{inter\text{-}res}$				$G_{GB\text{-}inter\text{-}res}$			
Mer [b]	Ile	Phe	Arg	Phe	Ile	Phe	Arg	Phe
1-in	-50.8	-71.8	-10.3	-75.6	6.6	17.0	-55.5	10.8
2-out	-57.3	-50.9	-138.4	-67.2	-1.8	-7.8	81.9	7.4
3-ext	-54.8	-51.9	-159.1	-72.1	3.1	-3.5	93.7	7.4
4-in	-55.9	-56.2	-176.1	-68.5	3.2	6.1	115.3	7.7
5-out	-45.3	-47.3	-172.2	-64.2	-3.3	-8.1	112.4	5.5

a) Inter-residue interactions are between the residue and the rest of the protein. $E_{inter\text{-}res}$ is the sum of inter-residue non-bonded and electrostatic energies. $G_{GB\text{-}inter\text{-}res}$ is the inter-residue GB solvent polarization energy. b) Rank of conformer after low temperature MCSA optimization runs starting from LEC of states 1* to 5* in Table 5.

Phe 288 and Arg 289 side chains of the computed DFP-like state (grey thin stick side chains label 2 in Figure 2) do not superimpose well on the 2DFP state (thick black stick side chains labeled 2DFP).

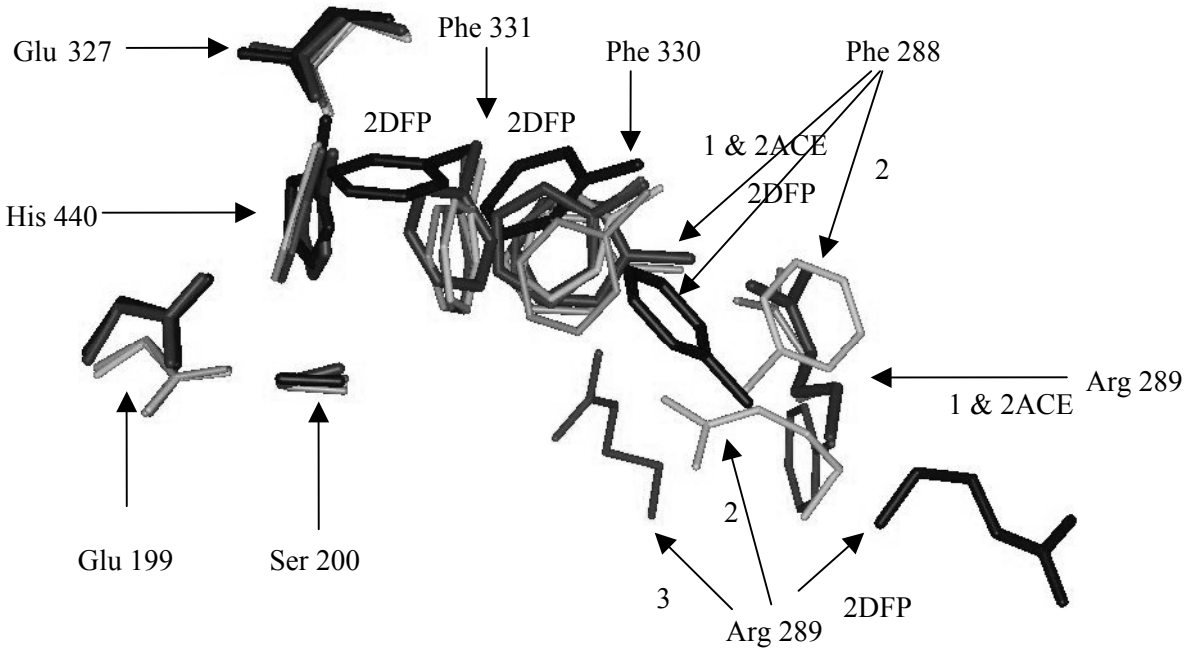

Figure 2. Gray scale stick model of catalytic side chains and select APL side chains in 2ACE and 2DFP (thick black stick) and the states ranked 1, 2 and 3 (thin grey stick), in Table VI. Ile 287 was left out for visualization purposes.

This difference is likely due to the fact that the 2DFP PDB structure has a covalently bound inhibitor in the active site. The computed ACE-like state (grey thin stick side chain label 1) superimposes very nicely on the native in-state (thick black stick side chains labeled 2ACE). In going from the lowest free energy state to the highest free energy state, the side chain of Phe 288 breaks pi-stacking interactions with

Phe 331 and moves completely out of the gorge (states 2 and 3). Arg 289 is outside the gorge in the 2ACE state and the computed ACE-like state. It is inside the gorge in the 2DFP state, the computed DFP-like state and the computed extended state.

CONCLUSIONS

The free energy of the ACE-like state is 4 kcal/mol less than that of the DFP-like state. Entropy favored the alternative states in the prediction calculations. A balance between solvation and the sum of electrostatic and non-bonded interactions governs the stability of the computed states. Solvation free energy favors the native state. The conformational energy favors non-native states. The differences in the physical properties of native and alternative states could be explained in terms of interactions with Arg 289.

ACKNOWLEDGEMENTS

Thanks to Frank Lebeda and Mike Lee for helpful comments. Sponsored by the Army HPC Research Center under the auspices of the Department of the Army. The content does not necessarily reflect the position or the policy of the government, and no official endorsement should be inferred.

REFERENCES

1. Raves ML, Harel M, Pang YP, Silman I, Kozikowski AP, Sussman JL. Nat Struct Biol 1997;4(1):57-.
2. Millard CB, Kryger G, Ordentlich A, Greenblatt HM, Harel M, Raves ML, Segall Y, Barak D, Shafferman A, Silman I, Sussman JL. Biochemistry. 1999 Jun 1;38(22):7032-9.
3. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE, The Protein Data Bank. Nucleic Acids Research 2000;28:235-242. http://www.rcsb.org/pdb/. Bernstein F, C, Koetzle TF, Williams GJ, Meyer EE, Jr. Brice MD, Rodgers JR, Kennard O, Shimanouchi T, Tasumi M, J Mol Biol 1977;112:535-.
4. Koellner G, Kryger G, Millard CB, Silman I, Sussman JL, Steiner T. J Mol Biol. 2000;18;296(2):713-.
5. Carlacci L Biopolymers 2001;58:359-373.
6. Carlacci L, J. Comput Aided Mol Design 2000;14:369-382.
7. Carlacci, L, Edison AS, Proteins 2000;40:367-377.
8. Qiu D, Shenkin PS, Hollinger FP, Still WC, J Phys Chem A 1997;101:3005-3014.
9. Dominy BN, Brooks CL, III, J Phys Chem B 1999, 103, 3765-3773.
10. IUPAC-IUB Commission on Biochemical Nomenclature, Biochemistry 1970;9:3471-.
11. Némethy, G.; Pottle, M. S.; Scheraga, H. A. J Phys Chem 1983;87:1883-1887.
12. Jorgensen WL, Tirado-Rives J, J Am Chem Soc1988;110:1657-.
13. Perrot G, Cheng B, Gibson KD, Vila J, Palmer KA, Nayeem A, Maigret B, Scheraga HA, J Comp Chem 1992;13:1-11.
14. Weiner SJ, Kollman PA, Nguyen DT, Case DA, J Comp Chem 1986;7:230-252.
15. Metropolis N, Rosenbluth AW, Rosenbluth MN, Teller AH, Teller E, J Chem Phys 1953;21:1087-92.